DOCUMENT RESUME

ED 340 767 TM 018 023

AUTHOR White, Edward M.
TITLE Assessing Higher Order Thinking and Communication Skills in College Graduates through Writing.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
PUB DATE Nov 91
NOTE 44p.; Commissioned paper prepared for a workshop on Assessing Higher Order Thinking & Communication Skills in College Graduates (Washington, DC, November 17-19, 1991), in support of National Education Goal V, Objective 5. For other workshop papers, see TM 018 009-024.
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Measurement; *Communication Skills; Critical Thinking; *Educational Assessment; Essay Tests; Evaluation Methods; Higher Education; Multiple Choice Tests; National Programs; *Portfolios (Background Materials); Problem Solving; Test Bias; *Testing Problems; Test Reliability; Test Validity; *Thinking Skills; Writing (Composition); *Writing Evaluation
IDENTIFIERS *National Education Goals 1990; Performance Based Evaluation

ABSTRACT

Writing encompasses a wide range of skills, from the mechanics of punctuation and spelling to the systematic or even creative development of ideas. The higher order skills of communication necessarily involve critical thinking and problem solving, and an assessment of the writing of advanced students will be able to encompass these other higher order skills. Three methods of assessing writing now exist: (1) multiple-choice tests; (2) essay examinations; and (3) portfolio assessments. The multiple-choice tests are not valid for many higher order skills, have a damaging effect on instruction, and are associated with racial and ethnic bias. The essay examinations have increased validity, though problematic reliability; are less damaging to instruction; and have been shown to be more fair to minorities. Portfolio assessment has the highest potential validity and the most positive effects on instruction. Although portfolio assessment poses problems in cost and reliability, it promises the most benefits to instruction and the most valid measurement of higher order skills. Any national assessment of writing should be principally or wholly a portfolio assessment. A 12-item list of references is included. Reviews by L. Boehm, J. Chaffee, and P. A. Facione of this position paper are included. (Author/SLD)

ED340767

Assessing Higher Order Thinking and Communication Skills In College Graduates Through Writing

Presented to the National Assessment Conference

Washington, DC: November 17-19, 1991

Edward M. White, California State University, San Bernardino

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ABSTRACT

Writing encompasses a wide range of skills, from the mechanics of punctuation and spelling to the systematic or even creative development of ideas. The high order skills of communication necessarily involve critical thinking and problem solving, and an assessment of the writing of advanced students will be able to encompass these other high order skills. Three methods of assessing writing now exist: multiple-choice tests, essay examinations, and portfolio assessments. The multiple-choice tests are not valid for many high order skills, have a damaging effect on instruction, and are associated with racial and ethnic bias. The essay examinations have increased validity, though problematic reliability, are much less damaging to instruction, and have been shown to be more fair to minorities. Portfolio assessment has the highest potential validity and the most positive effects on instruction; although portfolio assessment poses problems in cost and reliability, it promises the most benefits to instruction and the most valid measurement of high order skills. Any national assessment of writing should be principally or wholly a portfolio assessment.

TM018023

BEST COPY AVAILABLE

Problems of Definition

One of the root problems in assessing communications skills in writing is the wide variety of definitions, often unstated, that lead to different conceptions of what is being measured. These different definitions emerge from the complicated sequence of learning writing. The young child must learn such matters as the formation of letters, the system of writing from left to right on the line, and the spelling of words. Later on, the student will increase vocabulary, learn more complex sentence structures, create poems or stories, and begin to write about ideas. Still later, more advanced students construct paragraphs, essays, and research papers; learn to evaluate other peoples' ideas and incorporate them into their own arguments; and use writing as a central means of learning through discovery and criticism. All the earlier skills are necessary as part of the more developed skills.

Yet, as the learner advances, the skills do not merely develop; they take on a different shape and a wholly different source. The early skills are basically imitative: we must learn to spell and punctuate the way other people do, and we are taught the language forms of people we want to emulate. That is, the less sophisticated communication skills are a form of etiquette, based on the need to conform, and dependent upon socialization. But the more sophisticated skills depart from imitation and conformity. We do not want original spelling or

punctuation (though we often get them, at all levels) but we do want original thinking and independent critical problem-solving as part of higher order thinking skills. This difference between the imitative or socializing skills and the thoughtful or individualizing skills is crucial for thinking about assessment, for the imitative skills that may be appropriate to measure for younger children are likely to be inappropriate (or, at least, of minor importance) when we deal with higher order thinking skills. When we learn spelling and sentence structure, we do what we do specifically because that is the way they are done; that is, we do not think for ourselves. But when we develop arguments, conduct research, or solve problems, such imitation is not only insufficient, but it defeats the purpose; we must think for ourselves, as individuals, if we are to write well.

There is no question about the need for these individualizing skills as we consider writing as part of education for the work place or for citizenship. The information society of the future requires workers and citizens who have learned how to solve problems, to evaluate evidence, to come up with new ideas or new approaches to old ideas. As the SCANS Report for America 2000 (1991) puts it, "the problem-solving that people do at work requires complex, situated learning." Even more important, the basic premise of democracy demands citizens who do not merely do what they are told or think what they are told to think. Democratic political

theory rests on the premise that governmental authority depends on the "consent of the governed," and, as Thomas Jefferson argued, an uninformed consent is no consent at all.

Thus, for the purposes of this assessment, directed at the higher order thinking skills of college graduates, we need to be clear about just what we are measuring, and what we are not. The imitative and socializing skills may have some slight role to play, but our principal concern must be upon the ability of students to create discourse as a crucial part of the discovery, thinking, and evaluating process. We cannot rest content with a passive test of passive skills, but must come up with some way to assess actively the active thinking of students as they write.

Thus writing as an advanced skill becomes both the means and the expression of critical thinking and problem solving. The overlap is, of course, not perfect. Certain kinds of writing, such as poetry, journals, prayers, and the like move into realms of expression beyond the scope of the present project. And specialists in critical thinking and problem solving have come up with measures, often machine-scorable, that they claim will assess certain aspects of those mental processes. Some of those measures, if agreement can be reached about their value (no small task), may have a role to play supplementary to writing. But these newly-defined subdivisions of thinking have traditionally been part of rhetoric and logic, the ancient theoretical core of writing instruction; add grammar

and we have the medieval trivium, the core of all education a thousand years ago. Almost invariably, when critical thinking and problem solving have been taught, they have been taught through writing, most commonly these days in college courses in beginning or advanced composition. Thinking cannot occur in a vacuum, nor does sustained thinking result in answering machine-scorable questions. The world of work contains few machine-scorable tasks, but it is rife with writing assignments of all kinds. We cannot overlook the academic traditions of writing: essay exams, term papers, senior and post-graduate theses, books and articles, and so on, which for hundreds of years have led to and demonstrated learning and individualized thinking. Academic traditions and current practice confirm the central importance of writing for learning and demonstrating critical thinking and problem solving. And since we can assess the writing skills of students, we should not be content with less valid indirect measures of passive skills associated with some aspects of thinking. We should collect and assess the performance of students on complex writing tasks.

This active model of the important high-order communication skills requires an additional definition, that of thinking and writing as a process as well as a product. Indeed, when we look at writing in the context of critical thinking and problem solving, the writing process becomes far more important than the product. Since many teachers at all levels have now changed their teaching to reflect the process model, a national

assessment that supports improved teaching and enhanced student outcomes should certainly concern itself with more than writing products or correct answers. And the world of work, as well as the needs of democracy, must attend to how things are done more than to a brief sample of results. The process definition represents a substantial change in the way many teachers consider thinking and writing, and, since this change must affect the proposed assessment, we should consider its implications.

The process definition of writing has shifted the teaching of writing into the same context as the teaching of critical thinking and problem solving and somewhat diminished the role of writing as a mere agent of social stratification. As with any major shift in thinking, there are many causes of this new emphasis on the writing process: the onset of open enrollment in the City University of New York in 1970, research in sociolinguistics and dialectology demonstrating the value of variant dialects, heightened concern for racial bias in instruction and measurement as part of the drive toward a more racially open society, increasing worries about the place of creative thought in an increasingly passive school curriculum, fuller understanding by writing researchers of how writers actually produce texts, the increasing opportunities for employment of independent thinkers as well-paid rote jobs in industry disappear, and so on. Again the SCANS report reflects the need for process, in its definition of Thinking Skills:

"Thinks creatively, solves problems, makes decisions, visualizes, reasons, and knows how to learn." Whenever we consider writing as cause and demonstration of higher order thinking skills, we must define writing as more than a mere product to be graded for correctness.

We cannot ignore the evaluation of some writing products, however, despite the necessary emphasis on process; every sensible person knows that writing is, or can be, both a process and a product. We must recognize that a term paper, a letter of inquiry, a job application, an essay on a test, an article for publication, or a position paper such as this one is a product offered for evaluation and consumption. But the objection to defining writing as principally a product is nonetheless sound, for writing instruction and writing assessment. Writing on the job is more and more a matter of joint production, with teams of writers working together to produce reports, which are edited and published by a professional staff. The writer on the job must know how to produce drafts, and to revise work as a report develops. And almost all school writing is not inherently valuable or meaningful as product; in a world of waste paper, what is less likely to be published or even saved than old exams or essay assignments? Teachers do need to collect and grade occasional presentation copies of student work to be sure that students can, when needed, turn out presentable products. But teachers assign writing to students not so that they will create valuable documents but so that they will learn something

valuable while producing their usual mediocre documents. That is, school writing is necessary for instruction because the writing process itself is an important, probably the single most important, way to learn.

Writing, when defined as only a product, demands a high level of careful editing for correctness and neatness, which count more than such risky matters as originality, a personal voice, or even a developed idea. Teachers and tests, according to the product model, serve as society's gatekeepers, screening out those students whose written products do not measure up to the standards of school correctness and the school dialect; inevitably, such an evaluation heavily favors the privileged. This has been an important, if inglorious, function of college freshman English for over a hundred years. The first freshman placement test, established at Harvard in 1874, was designed, as James Berlin (1989) puts it with some irony, to ensure "that the new open university would not become too open, allowing the new immigrants, for example, to earn degrees in science or mathematics without demonstrating by their use of language that they belonged in the middle class" (23). We must be particularly careful to avoid perpetuating class, racial, and social distinctions in the name of writing assessment; the surest way to enforce such undemocratic class distinctions is to assess only (or principally) correctness of writing products, in the traditional way, according to the school dialect. We must assert clearly that correct writing is not the same as good

writing, despite pervasive social concerns for mechanical correctness.

The pernicious results of the traditional concern for mechanical correctness are apparent in most college students. One sign is the quaint faith in rules for writing, the belief that if one is sufficiently correct one will write well, a faith that is not only pervasive among old-fashioned teachers and test makers, but also firmly held by most students. Survey any group of college students and you will find that in their heart of hearts they believe that there are a handful of rules (which could be learned by sufficient drill) that would transform ordinary students into good writers. They, of course, don't know what these rules are, but English teachers are supposed to have the inside information and they aren't telling. If students could only learn these rules, they believe, they could then turn out perfect first drafts, which they believe good writers invariably do. In fact, every student has picked up and clings to one or two of these rules: never use the word "I," avoid all contractions, use lots of Latin terms such as "ibid," always use the comma after the name of a tree, and so on. The fact that the supposed rules are wrong or ridiculous in no way diminishes the fervor with which they are held, and the belief that a few more rules will make even those who cannot read into good writers forever never dies. Furthermore, they imagine that Miss Grundy handed out these rules in the fifth grade on the day they were sick; thus, everyone else <u>does</u> know all the rules, a belief

fraught with a subjective sense of frustrated inferiority in the face of now inaccessible knowledge.

I trace this painful perplex, identifying good writing as a product that follows supposed rules of correctness, to all the misplaced "grammar" instruction and reductive testing that students, particularly good students, receive on their way through schools dominated by the product model of knowledge in general and of writing in particular. While that instruction and testing hinders rather than helps their writing, it fills them with the belief in simple right answers about writing and thinking, and reinforces the destructive conviction that memorization rather than discovery is at the heart of writing and of knowledge itself. Students who envision writing in such a way tend to confuse plagiarism with research, passivity with seriousness, neatness with quality of prose, and fulfillment of the word count with fulfillment of the assignment. The thriving market in commercially or fraternally produced term papers is mute testimony to the destructiveness of this definition. As long as writing is defined only or principally as a finished product to be edited and corrected, writing remains a matter of guessing about the rules which govern correctness as defined by those in power. Such a definition is perfectly enforced by the usual multiple choice editing test.

Thus the process definition of writing, as well as the individualizing function of writing, should lie behind whatever assessment may evolve from the present quest. Process

evaluation argues for complex judgments of competence based on more than the correctness of the product. The process model sees writing as a series of overlapping activities, all of which have to do with critical thinking and problem solving: invention and prewriting, drafting, refining and rethinking, connecting, revising, and (finally) editing. The metacognitive activities associated with some definitions of critical thinking are an inherent part of the writing process, which requires that writers assess their work and their thinking in order to revise. The proposed assessment cannot rest content with passive testing of such matters as editing skills or neatness or even achieved organization of writing. Evaluation of writing as process needs to view drafts and revision, as clues to the thinking and writing that go on as the drafts become more and more finished.

Traditionalists suspect that this concern for process means abandoning standards, but proponents of the process model of thinking and writing argue that it moves standards from a rule book into the real world of work and of citizenship: this is writing seen as part of thinking, responding, evaluating, selecting and connecting ideas. Good grades become less a matter of memorization or of the dialect spoken in the home (traditionally, the nearer the home dialect to the school dialect the better the student's grade) than a measure of the thinking process of the student. The process model, by redefining reading and writing, does not abandon standards but

tries to redefine standards more substantially, more practically, and more fairly. No sensible teacher using the process model abandons the requirement for a well-edited presentation copy of student work, but the edited final draft is now just one part of the job, rather than the only thing that matters. Similarly, no sensible assessor of writing ignores wholly writing products, but a variety of products, with the revisions that led to the products, are now more important.

We must choose among a welter of theories and definitions in order to think clearly about what we are measuring when we assess writing and the thinking that it fosters and demonstrates. For higher order thinking skills, we must focus upon writing as a process, and writing as an active, individualizing skill intimately related to critical thinking and problem solving. While this is a complex definition, it responds to writing theory and current teaching of writing, as well as to the important place of writing on the job and for citizenship. Happily, recent developments in the measurement of writing ability allow us to come up with some procedures for arriving at such a complex and productive assessment.

Three Methods of Assessing Writing: Strengths and Weaknesses

The definition of writing that is assumed, as we might expect, leads naturally to a model of assessment. These models appear behind the tests that are often used to measure student performance. The simpler models, based on memorization and rules for correctness, lead to easy test-making and cheap scoring but

to a simplified result, a measurement unconcerned with active learning, the thinking and writing process, or higher order skills. The more complex models lead to more complex assessment, but with promise of valid measurement of what really matters. We should also remain aware that measurement affects curriculum, since teachers necessarily teach to tests; if the present effort leads to a writing assessment with a simplified definition and a reductive model, the assessment will degrade the teaching of writing and produce the reverse of what is intended.

We should also be aware that the nearest assessment in practice to the proposed assessment is the assortment of Graduate Writing Assessment Requirements (GWAR) in effect on many campuses and campus systems in America. Students must pass a GWAR to graduate from the California State University or City University of New York or State University System of Georgia, for example. Florida has its CLAST and Texas its TASP. Sometimes these high-stakes assessments screen students from advanced work; sometimes they serve as barriers to receiving the degree despite fulfilling all advanced work. Some of the assessments are tests, while some of them include alternate routes of demonstrating proficiency. Increasingly, portfolio assessment is working its way into these assessments. There is by now wide and deep experience in the assessment of writing, experience that should be referred to by those planning the new assessment.

There are three kinds of measurement now in use: multiple-choice tests, essay tests, and portfolio evaluations. I will briefly summarize their strengths and weaknesses for the current project.

Multiple-Choice Testing. The familiar multiple-choice tests are normally based on the definition of writing as product. Their assumptions are plain: (1) there are clear right and wrong answers to problems in the editing of texts, (2) the brief fragments of writing given on the test can elicit a student's knowledge of those answers, and (3) the student's ability to write (vaguely defined) will correlate highly with ability to edit correctly the prose on the test. As I have shown, the process and individualizing models of writing dispute each of these assumptions, and the definition of writing as a higher order thinking skill usually plays no role at all. Nonetheless, the multiple-choice tests continue to dominate, particularly in the schools, and wherever low cost is the primary concern of an evaluation program.

The advantages of this kind of test are familiar to everyone and result from the established traditional methods of devising a reliable norm-referenced test. Once the very high development costs (even higher than ever under truth-in-testing laws) are covered, the test can be easily administered and scored at low cost by computer. Once an appropriate (or even an inappropriate) norm group has been established, the norms can be published and score comparisons can be made. The statistics are

impressive and scoring reliability (though not test reliability) is essentially perfect. These advantages have fostered innovations in multiple choice testing, sometimes leading to machine-scorable tests that are designed to evoke more advanced abilities and more active performance. Many professional educators, administrators, and politicians are committed to these tests and the data produced by them, and the tests have a resounding appearance of objectivity. Some even call them "objective tests," as if they had been constructed by the same computers that score them.

These are powerful advantages for multiple-choice usage tests, powerful enough for them to survive and even to thrive despite the attacks of those who define writing by the process model. But if we do think of writing as a thinking and problem solving process, the disadvantages of those tests far outweigh their advantages. The difficulty is not alone that only a minor part of the construct "writing" is tested; the nature of the test itself, with its right "keys" and wrong "distractors," is opposed to the development of critical thinking. In more technical terms, the very narrow definition of writing as editing puts the validity of multiple-choice tests of writing in serious question despite their reliability. The validity issue becomes even more serious as evidence accumulates that the supposed correlation of scores on these tests with measures of actual writing is much higher for students who call themselves "white" than for those who are part of racial or ethnic

minorities (White and Thomas, 1981; Koenig and Mitchell, 1988). The argument that such tests are simply "indirect measures" and therefore harmlessly efficient pales before the simplicity of what most of them measure and the inevitable distortions of what teachers teach students to do as preparation for testing. As long as the validity and consequences of the multiple-choice tests of writing are so suspect, none of the other advantages should carry much weight.

Essay Testing. In response to the validity problem, we have seen major advances in essay testing over the last generation. Asking students to write on a writing test solves many, but by no means all of the problems I have just alluded to. Writing is still defined as a product to be measured, so the process model is still unsatisfied by such a test, but at least these tests measure a much wider range of writing activity than mere editing. Students need to demonstrate the ability to generate prose and to develop thought in prose, and so the measurement can respond to these parts of the writing process, as well as to editing in a more natural setting. An essay test can also ask for individual critical thinking and problem solving, and so can respond to a more sophisticated definition of writing. Most of the GWARs now in use depend heavily upon essay testing, which is generally supported by faculty and has credibility among students.

Of course, such a test is still an artificial situation for writing and still a small sample of the universe of writing;

nonetheless, such writing tests have become more and more widely used nationally for all kinds of purposes in recent years. They have become standard on college placement tests, for example, and on teacher certification skills tests, though they are often combined with other components to increase reliability. Many essay tests these days are carefully developed, using recognized procedures calling for test criteria and pre-testing of topics; rating reliabilities and costs (always the twin problems for writing tests) have become manageable using controlled holistic scoring (White, 1985).

The advantages of essay tests are clear: their increased validity, adequate reliability, and nearness to the process model of teaching have earned them support from teachers, many administrators, and an increasing number of test professionals. Since those being tested must generate actual writing, careful questions can lead to assessment of critical thinking and problem solving. The evidence of relative freedom from racial bias has been impressive. Furthermore, the necessary teacher involvement in creation and scoring has led to substantial in-service training of teachers as a useful by-product of the testing, and the mere existence of an essay test leads to an increase in the amount of writing that teachers ask their students to produce.

But the disadvantages of large-scale essay testing are also important. Even though development costs are much lower than for multiple-choice tests, scoring costs are higher. Although

it has become more and more customary to achieve very high inter-rater score correlations, it is difficult to achieve high overall test reliability with a single essay. If test reliabilities above .5 are needed, more than one writing sample must be obtained and scored. And test development is a never-ending process. The classroom effects of such testing, though they do enforce the need for students to write, are not wholly positive; many teachers will focus largely or even entirely upon writing under time constraints to set topics, thus sharply limiting the range of writing for their students. But the strongest objections to essay testing come from those committed to the process model, who are never altogether comfortable with tests that exclude revision, for them the most essential part of the writing process and an essential part of the construct being tested.

Portfolio Assessment. In response to these problems, portfolio evaluation has recently been adopted from the fine arts to writing measurement. For example, the State University of New York at Stony Brook has been for some years using portfolios to measure upper-division writing proficiency by re-evaluating the writing produced in lower division composition courses; Alverno College bases its advising and curriculum on individual student portfolios maintained over the full span of the college years; and public school systems in New York City, Pittsburgh, and the state of Vermont have adopted portfolios for proficiency assessment (DeWitt, 1991). Higher education, in

particular, has been most receptive to the concept. One estimate declares that fully one-third of American colleges and universities are experimenting with portfolio assessment for various purposes as we move through the 1990s.

There are good reasons for this enthusiasm for what seems to be an exciting new idea in writing assessment. Portfolios have shown themselves to be valuable for performance evaluation in the arts for many years and seem just as appropriate for the evaluation of writing. They offer many new possibilities for evaluating the writing process, as well as for assessing writing and general education programs. Portfolios use documents already produced, instead of generating new ones, and they can make a stronger claim for validity than can any other assessment, since they include "real writing" not generated under testing conditions. They can give us, to use the American Association of Higher Education's Pat Hutching's memorable metaphor, a motion picture of student performance, as opposed to the snapshot given by a test.

But we need to temper this enthusiasm (which reminds me of similar passion on the part of essay testing advocates twenty years ago) with some caution.

When we speak of portfolio assessment, we are really speaking of two separate procedures, each of which has its own problems. The first issue has to do with the collection of materials for assessment. The second set of issues, which has received rather less attention, has to do with assessing the

materials after the collection has been made. Each of these issues contains a host of problems which need to be dealt with.

I will list here just a few of those problems, which I have no space to discuss: What is to be included in the portfolio? Who will be responsible for collection and verification of materials? How is the sheer bulk to be managed? What kind of assessment is possible and affordable? How can the assessment be fair for transfer students or for those with heavy outside demands on their time? How can an assessment at the upper-division college level be fair to students whose major professors assign little or no writing, or who give assignments that require formula or simply bad writing? Should the original instructor's grades and comments be left on the papers in the portfolio, or removed? How can reliability and validity of measurement be demonstrated?

But it is clear that numbers of portfolio experiments have begun to come up with answers to these problems. The largest experiment at the college level, to my knowledge, is the portfolio placement program now in effect (with funding from the Fund for the Improvement of Postsecondary Education) at Miami University of Ohio; the portfolios there consist of four specified pieces of writing, authenticated by a teacher's signature. The portfolios are scored accurately by readers, using a scoring guide, at the rate of seven minutes per grade; more than three hundred portfolios are scored at a rating session, at modest cost and with high reliability. Similar

reports of accuracy and cost-effectiveness are starting to emerge from many campuses. (Pat Hutchings has been accumulating information on most of these assessments and is a ready source of documentation.) Portfolio assessment of writing has by now proven itself to be the most instructionally supportive procedure available at this time, and it is clear that those working with portfolios are developing procedures and practices which will make it increasingly practical.

Thus, portfolio assessment is the obvious wave of the future, the most responsive of all assessments to the process model of writing and to writing defined as a higher order thinking skill. Depending upon how the contents of a portfolio are defined, it can include drafts as well as presentation copies, and writing across many courses and years of study. It incorporates assessment productively into teaching and learning, demonstrates that learning and recording learning are the responsibility of the student, and has a most constructive effect upon what teachers do in the classroom. Any large-scale effort to develop a new model of writing assessment ought to devote itself to ways of using portfolios as a major measurement device.

It is true that the problems in using portfolios for large-scale assessment and data collection look daunting, but some reflection shows that these problems are not nearly as difficult as their sheer novelty makes them appear to be. Huge savings in assessment costs result from evaluating existing

student materials: there is no expense for test development or test administration. Furthermore, the virtually insuperable problem of motivating advanced students or college graduates to take an assessment of no perceived value to them now disappears; the materials have already been produced under the standard motivations that universities provide. Instead of spending vast sums of money and vast energy on test development and test administration, and even more of both in fruitless efforts to induce unwilling adults to take seriously one more test, a portfolio assessment can spend its money and time in improving the collection and scoring of existing materials.

Such an assessment can build on the reservoir of experience and goodwill now accumulating around the portfolio experiments on the most innovative campuses in the country. And a national portfolio assessment could count on the support of faculties, since it shows respect and support for their work rather than the suspicion inevitably expressed by a totally new test. Such support is crucial, if the assessment is to reach its goal of improved student performance. Measurement by itself will not accomplish that goal, particularly if the results of the measurement are felt to be at odds with the most important goals and results of instruction. If the goals are to be reached, universities and their faculties will need to believe the measurement and be part of it.

Recommendations

What Ought Not to be Done. Since assessment has often been reductive in its approach to writing and irrelevant or even destructive in its approach to the teaching of writing, the new assessment should deliberately and consciously seek to avoid repeating those patterns. Specifically, the assessment should not do the following:

1. A new national assessment should not merely repeat the most convenient and economical of present assessment procedures. The federal government should learn from what has gone wrong with many large-scale assessments in the past. In particular, any national assessment should be fully aware that reductive large-scale tests, given to students of widely different backgrounds and locations, have not supported improvement of instruction or learning. We should not administer an assessment that merely imitates the tests--and the mistakes--of the past.

2. We should not begin an assessment without considering carefully what kind of results we seek. If we are looking for confirmation that students from secure and privileged families do better on tests than students with insecure and deprived families, we will assuredly produce such results; but they will tell us nothing new and be of no use. On the other hand, if we are looking for an assessment procedure that will support the teaching and learning of writing as a high order skill, that will help schools and teachers and students see how they are

doing on what really matters, and that will promote the goals of a truly literate democratic society, we will develop an assessment to reach those goals. This suggests that we should not imagine that critical thinking, problem solving, and communication can be separated into measurable pieces; we should seek to assess all three through their most powerful demonstration in writing. This means that we should not simply administer a reductive multiple-choice or essay test, yielding simple numbers that can be used only for political purposes.

3. We should not ignore the vast differences in dialect, privilege, geography, culture, and almost everything else that appear in the United States. That is, we need to follow the recommendation of the Nation Council on Education Standards and Testing (1991) that "endorsed a national assessment system over a single national test to measure student achievement." If we confuse national achievement standards, which can be set in a useful way, with a national test, which cannot, we make a fundamental error, confusing goals with procedures. Any useful assessment system must recognize regional, or even local, assessment responsibilities.

4. Finally, the national assessment cannot be wholly top-down if it is to be effective. It must include the active participation of teachers, as well as the support of the local college community, if it is to lead to the improvements everyone seeks. Any assessment that is perceived to be external to the interests and values of a particular school will be an

irrelevant annoyance rather than a spur to achievement in the school.

What Ought to be Done. The national assessment should assume the following responsibilities:

1. Establish clear definitions of writing as a high order thinking skill, demonstrating critical thinking and problem solving, and most importantly manifested as a process rather than merely a product.

2. Recognize that some form of portfolio assessment (which may include in the portfolios some timed essay test writing or other responsible testing) is the most effective way to meet the need for national standards and a national assessment system.

3. Convene a conference of those most actively involved in portfolio assessment at the college level, particularly those administering funded projects, to focus attention on the ways in which the problems of portfolio assessment have been and may be handled. The new national assessment should become a center of information on portfolio assessment.

4. Develop a series of descriptive national standards for writing at the freshman, sophomore, junior, and senior college levels. These descriptions should not only take into account the essential place of writing for critical thinking and problem solving, but also recognize that more advanced college students have entered the discourse communities of their major fields of

study. Senior portfolios of physics majors, for example, should demonstrate the ways in which the physics community poses questions and proceeds to deal with them; political science or art majors will not follow the physics pattern, but rather use the languages and procedures of their fields.

5. Develop a series of sample portfolios, illustrating national standards of writing performance at various levels, that can be met in various ways. Thus, a statement about what college freshmen should be able to do might be supported by a series of portfolios exemplifying those standards, or illustrating different levels of performance in relation to those standards. The sample portfolios for graduating seniors will need to show how the large variety of major fields of study can meet national standards; the portfolios should be exemplary rather than inclusive or conclusive.

6. Foster experimental assessments in some key communities or states that will use the national standards and sample portfolios as guides. If due care and consultation have led to responsible standards and useful sample portfolios, the validity of the assessment should not be an immediate problem. Methods for demonstrating reliability will need to be developed, so that local assessments can be depended upon and related to the national standards. Widespread participation in these experimental assessments will lead to widespread acceptance of the meaning and importance of the results.

7. Disseminate results of the experimental assessments and

demonstrate the value of the assessment process, as well as of its products. Feature assessments which have led to enhanced performance and provide a variety of models to the higher education community. Any comparison of institutions should take into account the wide disparity of student populations in American higher education. Reports should avoid simple and misleading numerical data, but be rich in examples of different kinds of portfolios from different kinds of institutions.

8. Develop the national assessment system as a support center for encouraging local assessments, gathering and disseminating information, maintaining records, and fostering standards. Through these activities, the assessment will actively move the nation's students and institutions towards the fulfillment of the goals. At the same time, the assessment will be able to document the progress of the country toward national goals by reporting the percentages of different populations that meet the established standards as each year passes.

Conclusion

Assessment is much more than testing, and a national assessment worthy of the name will do much more than impose a national test on an unwilling nation. An effective national assessment will foster and support regional assessments through portfolios and other process measures, in order to help students develop the high order thinking and problem solving skills that the national interest demands. Through centralized funding,

information and data gathering; by disseminating exemplary procedures and results; and by maintaining clear goals and standards, a national assessment can elicit from states and localities the participative assessments that lead to genuine and measurable improvement. We should expect improvement in performance as a result of the assessment, which must not be allowed merely to report what everyone knows is amiss. Such an assessment will require time, much consultation, and imaginative solutions to the problems of reliability and funding. But the power of this kind of assessment system to improve American education is immense. The present opportunity to turn assessment in this positive direction offers significant hope for the future.

San Bernardino, California, September 20, 1991.
Revised November 22, 1991.

References

Adelman, C. (Ed.). (1988). Performance and Judgment: Essays on Principles and Practice in the Assessment of College Student Learning. Washington: U.S. Department of Education, Office of Educational Research and Improvement.

Belanoff, Pat, & Marcia Dickson. (1991). Portfolios: Process

and Product. Upper Montclair, N.J.: Boynton/Cook.

Berlin, J. (1989). Rhetoric and Reality. Carbondale, Ill.: Southern Illinois University Press.

De Witt, K. (1991). "In Vermont Schools, Test on How Well Students Think Draws New Interest." The New York Times, Sunday, September 1.

Elbow, Peter, & Pat Belanoff. (1986). "Portfolios as a Substitute for Proficiency Examinations." College Composition and Communication 37 (1986): 336-39.

Koenig, J. & Mitchell, K. (1988). An Interim Report on the MCAT Essay Pilot Project. Journal of Medical Education, 63, 21-29.

Mitchell, K., & Anderson, J. (1986). Reliability of Holistic Scoring for the MCAT Essay. Educational and Psychological Measurement, 46, 771-775.

National Council on Education Standards and Testing. (1991). "Standards, Testing Council Endorses Assessment System." Education Daily, August 16.

Ruth, L. & Murphy, S. (1988). Designing Writing Tasks for the As

sessment of Writing. Norwood, N. J.: Ablex.

Secretary's Commission on Achieving Necessary Skills. (1991). "What Work Requires of Schools: A Scans Report for America 20000." U.S. Department of Labor.

White, E. M. (1985). Teaching and Assessing Writing: Recent Advances in Understanding, Evaluating, and Improving Student Performance. San Francisco: Jossey-Bass.

White, E., and Thomas, L. (1981). "Racial Minorities and Writing Skills Assessment in The California State University and Colleges." College English, 42, 276-283.

A Review of Edward M. White's "Assessing Higher Order Thinking and Communication Skills in College Graduates"

Lorenz Boehm

Oakton Community College

I like and agree with the central ideas of Professor White's essay. Based on my own reading, research, and experience, I'll say the claims he makes for writing assessment are both sound and valid, particularly his claims (a) that seeing writing as a process is central, (b) that "the writing process itself is . . . probably the single most important way to learn," and (c) that the process of writing is an especially, even a singularly, effective way of assessing higher order thinking skills.

I also agree with his recommendations, particularly -- (a) that machine scored multiple choice tests of grammar, punctuation and spelling are conceptually mistaken, typically assess exactly the wrong things, undermine attempts to teach both writing and critical thinking, and should have no role to play in any assessment of critical thinking; (b) that having students write an essay is generally a better assessment of their writing and thinking abilities than is a multiple-choice test; and (c) that, by far, the very best way to assess writing and critical thinking is through portfolios. That last is gutsy; it's also, in my opinion, ideal.

Seeing the value of writing as a process is the key -- because of what it lets us see or know about a mind and how it works, or doesn't. For a good while now, writing teachers, and, increasingly, teachers in other disciplines, have maintained that most writing, and especially essay writing, unavoidably, involves critical thinking. They argue that to write essays is to think critically, that the process of writing is critical thinking. Look at what's involved in that process: selecting a focus, finding a way to hold it still, generating ideas about it, choosing from among them, choosing words, ordering them, rejecting, keeping, reconsidering, polishing. Unpack composing an essay, unpack almost any college-level writing assignment, even, although there will be less there to unpack, the well-known, "canned," first-semester-freshmen, five paragraph essay; certainly unpack an exploratory piece where the student is exploring new ground -- look at what mental work doing it requires -- and you find it involves taking things apart, seeing relationships, making connections, using judgement, reasoning, evaluating, inferring, and so on -- elements, each of them, surely, in any credible definition of critical thinking.

It's an important point, substantiated by a considerable amount of the scholarship generated by the writing across the curriculum movement over the past fifteen years.[1]

Thus it follows, from this point of view, that writing is a perfect medium, perhaps the perfect medium, for assessing thinking. Assess the writing process, and you assess the

thinking process; they are inseparable.

Further, arguably, no matter how critical thinking is defined, writing remains an ideal medium for both teaching and assessing it. Professor White does not offer a particular definition of critical thinking, let alone a set of "sub-skills"; instead, he uses a number of different definitions -- including "original thinking," "independent critical problem solving," "discovery," "coming up with new ideas," and the ability to "develop arguments," "do research," and "evaluate evidence." Or he simply quotes the SCANS Report's definition: "thinks creatively, solves problems, makes decisions, visualizes, reasons, and knows how to learn." I'm not especially troubled by that; his point, I believe, is that, no matter how critical thinking is defined, essays in portfolios are the best way to assess it.

As an alternative, we can consider, if only for a moment, what has been, very likely, the most often quoted definition of critical thinking, that written by Robert Ennis:

> Critical thinking is reasonable, reflective thinking that is focused on deciding what to believe and do.[2]

Add Ennis' fourteen "Dispositions" and twelve "Abilities," which range from "seek reasons" and "be open-minded" to "making value judgements" and "defining terms" and which include as tidy and well-thought-out a set of "sub-skills" as we're like to find this side of New Jersey. No problem. At every level of instruction, countless writing teachers spend endless hours designing writing activities engineered to do or facilitate exactly that, or them.

And the conviction that writing can do it resides happily in the marrow of their professional bones.

There are, of course, other definitions. I'm using Professor Ennis's here only to show the compatibility of writing with assessing one particular (and particularly well-developed) definition of critical thinking. It represents one point of view, or "camp" in the critical thinking movement. Another, and one I much prefer, is represented by Mark Weinstein, Co-director of the Institute for Critical Thinking at Montclair State, who defines critical thinking as doing the mental work of the academic disciplines. Says Weinstein:

> Critical thinking in the disciplines (at the undergraduate level) requires mastery of the forms of inquiry. Embedded in language, such forms yield the tools for inventing, organizing, communicating, and utilizing the content of the various areas of human concern.

A radically different definition, but, no matter. Writing still works at both teaching and assessing it. All that changes is the number of faculty members using it; not the belief in its role. Says Toby Fulwiler: "writing progresses as an act of discovery; no other thinking process helps us so completely develop a line of inquiry or mode of thought." And the writing-as-process people, I among them, nod sagely, smile brightly, and think of portfolios. And rightly so. Given the range of writing/thinking a portfolio can accommodate, and given its ability to reflect the stages a piece of writing/thinking has gone through, it's hard to imagine a better assessment instrument.

Thus, I have few concerns about the rightness of Professor White's claims and recommendations. However, I do have some concerns, of another sort.

Once again: seeing the value of writing as a process, a thinking, making-meaning process, is the key. Those who see its value are believers, devotees; they use it at every turn. They swear by it. However, those who don't see it's value -- don't use it. They swear at it. And, while this includes a lot of teachers to be sure, it also, unfortunately, includes a lot of department chairs; deans; vice presidents of curriculum, instructional support, and finance; college presidents, and members of boards of trustees. And assessment test designers. Thus, my concern.

Professor White, having argued, rightly, in my opinion, that portfolios should be, or should be part of, the instrument of choice for a national assessment of critical thinking, then acknowledges that there are "a host of problems which need to be dealt with before that instrument can be put in place." And he delineates, as he says, "just a few" of them. It's an astute list; clearly he has walked the turf more than once. And he is confident the problems can be solved. He is a believer, a devotee; he swears by, not at, the writing process.

What Professor White sees and calls a collection of questions and problems that surely can be answered and solved, others see as a delineation of all the reasons why portfolio assessment won't work. The simple fact that there are so many

"problems" argues against it: it's too bulky, too unwieldy, too unmanageable, too expensive; it requires too many committees; it's too unpredictable; there are too many unkowns. The unbelievers speak.

I don't mean to be glib (or fresh). My experience working with community college faculty and administrators over the past fifteen years on writing across the curriculum programs or on writing assessment indicates that, while there are times when the "problems" at a school simply can't be solved, almost always what gets in the way is attitude. The believers make it work; the unbelievers make excuses.

Because I like and agree with Professor White's claims for writing as a process (I am, afterall, a believer), I also like and agree with his recommendations. I too want to say, "...portfolio assessment is the obvious wave of the future, the most responsive of all assessments to the process model of writing, and to writing defined as a higher order thinking skill." (p. ~~26~~ 21) Yet, my hunch is, in order to make that work, the national assessment administrative structure has to include a mechanism for bringing faculty and administrators into the fold; they too will need to see, believe in, and value the possibilities of the writing process. Yes, educators have to "recognize that some form of portfolio assessment. . . is the most ~~responsible~~ effective way to meet the need for national standards and a national assessment system" (p. ~~21~~ 25), but I don't believe we can't legislate that recognition. (We <u>can</u> legislate machine-scored,

multiple-choice tests; they are more asily managed and are far less hassle.) In order for national assessment of critical thinking to be done through writing only, let alone a portfolio of writings, we are going to have to bring people to it. Make them believe, or else it will not work.

Endnotes

1. As I indicated, the literature on the relationship between writing and thinking is extensive. I would list the following as among the most useful:

   John Bean , et al, "Microtheme Strategies for Developing Cognitive Skills," in New Directions for Teaching and Learning, C. W. Griffin, ed., Jossey-Bass, 1982.

   Ann E. Berthoff, "Towards a Pedagogy of Knowing," Freshman English News, Spring, 1978.

   Janet Emig, "Writing as a Mode of Learning," College Composition and Communication, March, 1977.

   Toby Fulwiler, "Writing: An Act of Cognition," in New Directions for Teaching and Learning. See above.

   J. N. Hayes, et al., The Writier's Mind: Writing as a Mode of Learning, National Council for Teachers of English, Urbana, Il, 1983.

   Judith A. Langer and Arthur N. Applebee, How Writing Shapes Thinking, NCTE Tesearch Report No. 22, 1987.

2. This definition is available in a variety of places, but because readers may well be interested in other essays in the collection, I'm quoting here from Robert Ennis, "A Taxonomy of Critical Thinking Dispositions and Abilities," in Teaching Thinking Skills: Theory and Practice, edited by Joan Boykoff Baron and Robert J. Sternberg, W. H. Freeman and Co., 1986.

Fiorello H. LaGuardia Community College THE CITY UNIVERSITY OF NEW YORK
31-10 THOMSON AVENUE, LONG ISLAND CITY, N.Y. 11101 • Telephone (718) 482-7200

John Chaffee, Ph.D.
Director, Creative and Critical Thinking Studies
LaGuardia community College; The City University of New York

Review of the position paper: *Assessing Higher Order Thinking and Communication Skills in college Graduates*
Edward M. White, California State University, San Bernardino

1. What abilities (critical thinking, communication, problem-solving) have been identified and why were they selected?

This paper focuses exclusively on written communication. The author contrasts writing as an imitative, mechanical skill with writing as an "individualizing," thoughtful endeavor. It is this latter conception of writing which should be the concern of the NCES project: "...our principal concern must be upon the ability of students to create discourse as a crucial part of the discovery, thinking, and evaluating process." This general orientation is certainly appropriate for this project, although the author does not articulate in any more detail the specific language/thinking abilities which characterize this "higher-order" writing.

2. Are the selected abilities appropriate in meeting the goals of this project?

As noted, "higher-order" writing abilities should certainly constitute an essential component of the NCES assessment, although these abilities would naturally have to be clearly defined and their relation to "higher-order" thinking abilities established.

3. Are the abilities defined in a way that would make possible assessing their development in college students?

The author makes a compelling case for assessing the *process* of students' writing as well as their writing *products*. "Thus the process definition of writing, as well as the individualizing function of writing, should lie behind whatever assessment may evolve

from the present quest. Process evaluation argues for complex judgments of competence based on more that the correctness of the product. The process model sees writing as a series of overlapping activities, all of which have to do with critical thinking and problem solving: invention and prewriting, drafting, refining and rethinking, connecting, revising, and (finally) editing. The assessor cannot rest content with passive testing of such matters as editing skills or neatness or even achieved organization of writing. Evaluation of writing as a process needs to view drafts and and revision, as clues to the thinking and writing that go on as the drafts become more and more finished."

However, beyond this general description, the author does not define the abilities to be assessed.

4. Do the proposed assessment methods allow for: accurately measuring the abilities; determining the acquisition barriers; identifying effective learning environments?

The author evaluates three methods for evaluating higher-order writing abilities: *Multiple-Choice Testing* (unsuitable); *Essay Testing* (more suitable, but still limited to the product of students' writing); *Portfolio Assessment* (most suitable -- permits assessment of students' writing products and processes). In theory, portfolio assessment should be responsive to these criteria.

5. Are the methods or suggestions presented practical, replicable and complete?

Portfolio assessment is an emerging methodology. The complexity of this approach raises serious questions regarding its feasibility for a project such as this. Perhaps some of the key elements of this approach might be integrated with a more traditional essay test -- for example, giving students the opportunity to revise their essay in response to new information or some form of feedback.

6. General Comments:

Conclusion: This paper underscores the crucial relation between critical thinking and critical literacy, a relationship that should be incorporated into any assessment of cognitive and communication abilities.

REVISED Nov. 25, 1991

A Critique of Edward M. White's

"Assessing Higher Order Thinking and Communication Skills"

Prepared by Peter A. Facione

Dr. White's paper forcefully advocates a portfolio based system for purposes of assessing writing as process, not simply product. Dr. White argues that any strategy using standardize objective testing or essay testing is unredeemably flawed because necessarily it will be socially biased, lacking in construct validity, artificial, or fail in principle to capture writing as process. Portfolio assessment, which he calls "the obvious wave of the future," is claimed to be the only method capable of targeting writing as process and hence the only valid assessment strategy.

To bolster his position, Dr. White mentions "experiments" occurring around the country which have "begun to come up with answers" to a host of daunting problems such as reliability, validity, cost, physical bulk, materials verification, what to include, what to exclude, who will make the actual assessments, and how to adapt the portfolio strategy for use beyond the area of the visual and performing arts.

Threading its way throughout Dr. White's paper is the often repeated proclamation that higher order thinking or critical thinking simply cannot be assessed in any reasonable way other than by the portfolio strategy. And, since critical thinking is essential for good writing, the argument for process oriented portfolio assessment is all the stronger.

Dr. White's position, however, is unconvincing for many reasons. First, it rests on too many unproven empirical assumptions, not the least of which is that there is no valid way to assess a process other than by watching the process as it happens. We don't have to watch healing (or water coming to a boil, or learning) to determine at some later time if healing (or boiling, or learning) has occurred. If Dr. White's assumption were actually true, then even portfolio assessment would also not work. The drafts he proposes that we might gather remain only snapshots, an incomplete photo album at best. But the fact is it may be possible to use other strategies, including pretesting and posttesting to see the effects of a process and establish developmental norms.

Second, in his rush to embrace the portfolio strategy, Dr. White gives neither standardized testing nor essay testing a fair hearing. The field of educational testing and measurement does not have to be reinvented. There is a considerable literature which documents the strengths and weakness of each strategy. There are counter-examples to Dr. White's overstated criticisms and ways to overcome the problems his paper cites.

For example, contrary to Dr. White's assertion, standardized tests have been shown

to be valid measures of critical thinking. ETS addresses effectively addresses higher order thinking skills on the GRE, the LSAT and some parts of the SAT. The newly published California Critical Thinking Skills Test (CCTST) has demonstrated content validity, construct validity, reliability, and is free from gender bias and ethnic bias. The CBEST (California Basic Educational Skills Test), addresses a similar set of skills but at a more primitive level. The research supporting the CCTST illustrates that Dr. White's claims about not being able in principle to test core college level critical thinking skills using objective, standardized testing are simply wrong.

As a general assessment strategy standardized, objective multiple-choice testing (or short answer, or multiple-rating item testing, which are more problematic variations on the same thing) offers the advantages of concurrent validity, inter-form reliability, inter-rater reliability, and statistical precision to a degree which would permit judgements along a significant continuum (percentile range). Aside from being substantially more efficient, these clear testing advantages, should have been taken into account. Many of these same testing advantages, except for degree of precision due to weaker inter-rater reliability, can also be achieved by essay testing.

In contrast to portfolio assessment, essay testing at least starts from a common point of reference for all those taking the test -- the well conceived question prompt -- and it still suffers severely from problems of cost and weak reliability. Portfolio assessment, which is more labor intensive and which permits vastly greater variability in the materials to be evaluated, will be hamstrung by these twin problems. Experiments in college departments with portfolio assessment of their own students certainly can be made to work. But they do not make the case that the portfolio strategy is the national assessment panacea.

One version of what Dr. White is suggesting is that throughout the nation people be allowed to submit for assessment a collection of their writings, preferably including drafts. For example, among other things submitted, a 23 year old white liberal arts student from Connecticut might turn in the third draft of a term paper on Plato, an 18 year old hispanic high school graduate from Los Angeles a poem written for a sibling's wedding, and a 36 year old single parent a practice memorandum prepared in a night school office management class..., well give us a break. Where is the validity? Where is the reliability? Where is the writing process? Where is the fairness?

The portfolio strategy works best when the persons being assessed come from a common background (e.g. all were once music majors) or from the same program (e.g. all are currently enrolled in an international studies capstone seminar course) and when all the judges who jury the portfolios have well developed and commonly accepted standards (e.g. all have expertise in the area to be assessed, all agree to use the same grading rubric and criteria, and all been trained and have practiced on sample cases to refine their inter-rater reliability coefficient.) Even if all the other problems are resolved, and many are much more than the technical glitches Dr. White's paper seems to imply, these two necessary conditions will not be able to be met.

The question of which testing strategy to use is more empirical than philosophical. It is not true that objective testing or essay testing in principle must fail.

If anything, given the diversity among those to be assessed and among those doing the proposed portfolio assessment across the entire nation, the portfolio strategy appears to be a very unwise choice. This raises a serious public policy issue. How will this strategy ensure a sufficient level of consistency of judgement to assure the public that the national assessment process is valid, reliable, and fair?

Where might we look for a strategy suitable for the assessment of writing? The most promising approach appears to be a combination of objective, standardized tests and essays. The essays provide the writing samples, two if necessary, and the standardized tests target the critical thinking processes and the basics of putting together understandable sentences. Will this capture all of good writing? No. Will it allow the full range of creative thinking and expressive prose to emerge? No. In principle, can it work as a way of assessing college level critical thinking and writing for purposes of the productive and effective participation in American economic and political system? Yes.

ETS in Princeton NJ can provide basic information about the LSAT, GRE, and SAT. The American College Testing Service can provide information about the Advanced Placement program in composition. Other national testing services and regional educational research labs are useful sources of additional information. The thousands of publications on student assessment, program evaluation, and educational testing are so vast and varied that there is an entire ERIC Clearinghouse devoted exclusively to this concern.

The four Technical Reports supporting validity and reliability of the California Critical Thinking Skills Test, published in 1990, are ERIC documents: ED 327 548 "Experimental Validation and Content Validity," ED 327 550 "Factors Predictive of Critical Thinking Skills," ED 326 584 "Gender, Ethnicity, Major, CT Self-Esteem, and the California Critical Thinking Skills Test, and ED 327 566 "Interpreting the CCTST, Group Norms, and Sub-Scores".

Please Note: Dr. Facione's comments were issued from an earlier version of Dr. writer's paper. Therefore, ~~same~~ Dr. White may have addressed some concerns in this revised version.